Application Performance On Multicores

Douglas Doerfler

CSRI Seminar

Feb. 4th, 2008

SAND2008-1084P

Unlimited Release

Printed February, 2008

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

NNSA National Nuclear Security Administration

Sandia National Laboratories

Executive Summary

It’s the memory subsystem!
(not quite, but pretty darn close)

What my talk is NOT about

- Massively parallel performance (I.e. performance off the node)
- Software programming models and Operating System Impact
  - A topic of a future System Engineering Seminar Series (SESS)
- Architecture research

Multicore is here, are we ready?

- Dual-core is mainstream
  - Not a big deal, process level parallelism
- Quad-core is almost mainstream
  - In general, still no strategy for SW
  - It's an SMP?
  - MPI everywhere?
- Eight core is in the near future
- 10's to 100's of cores in the next 3 to 5 years?
- Platform Roadmap
  - Red Storm - dual-core AMD
  - TLCC - quad-core AMD
  - ASC/NMHPC next generation capability system - N-core ?
  - ASC/Sequoia UQ platform in 2010 - ?

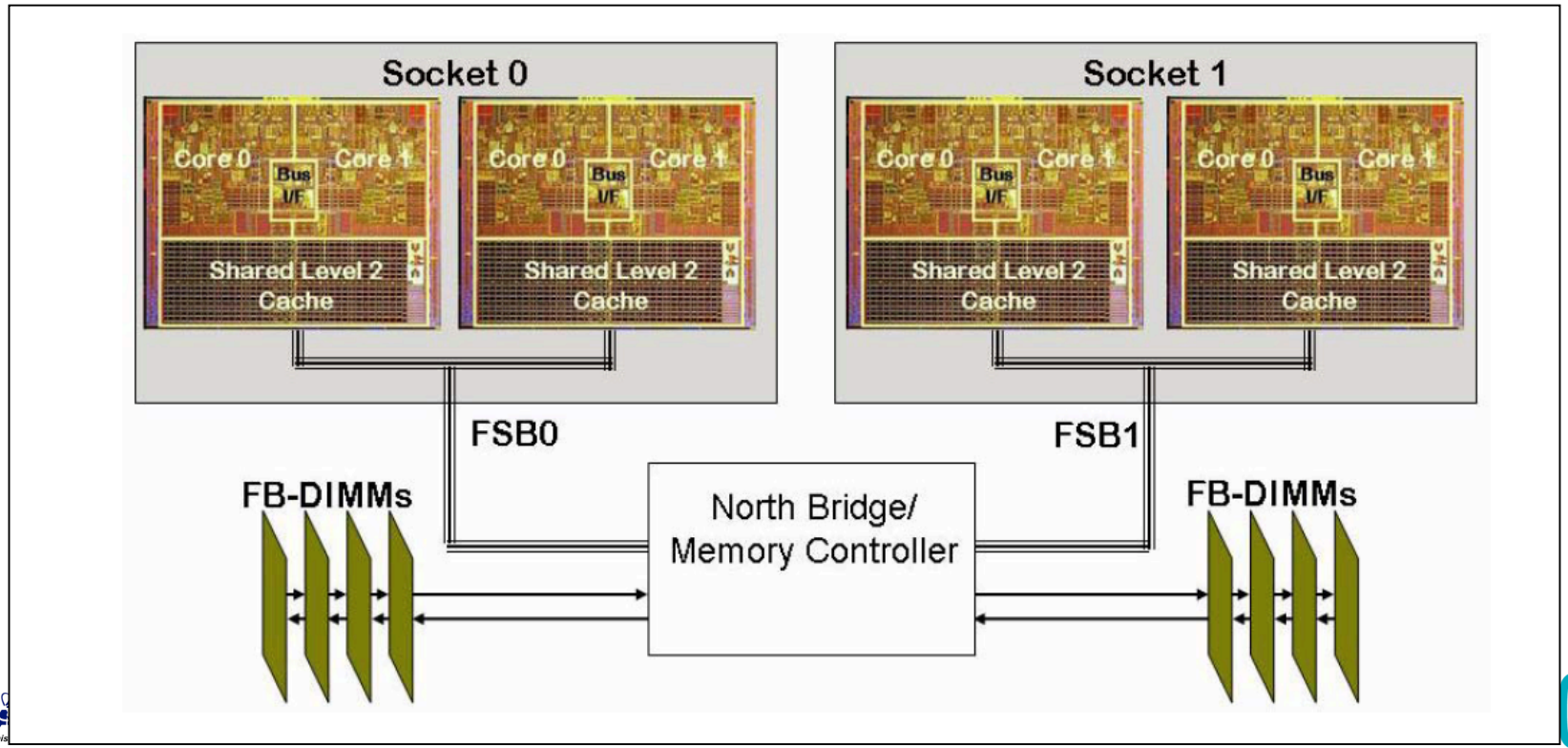
What is Multicore?

- For this talk, general purpose processors only
- Multicore comes in many flavors
  - How does a processor company differentiate?
  - Distinguishing architecture features
    - Cache hierarchy
    - Bus
    - Memory controller
    - Core architecture, including # of cores

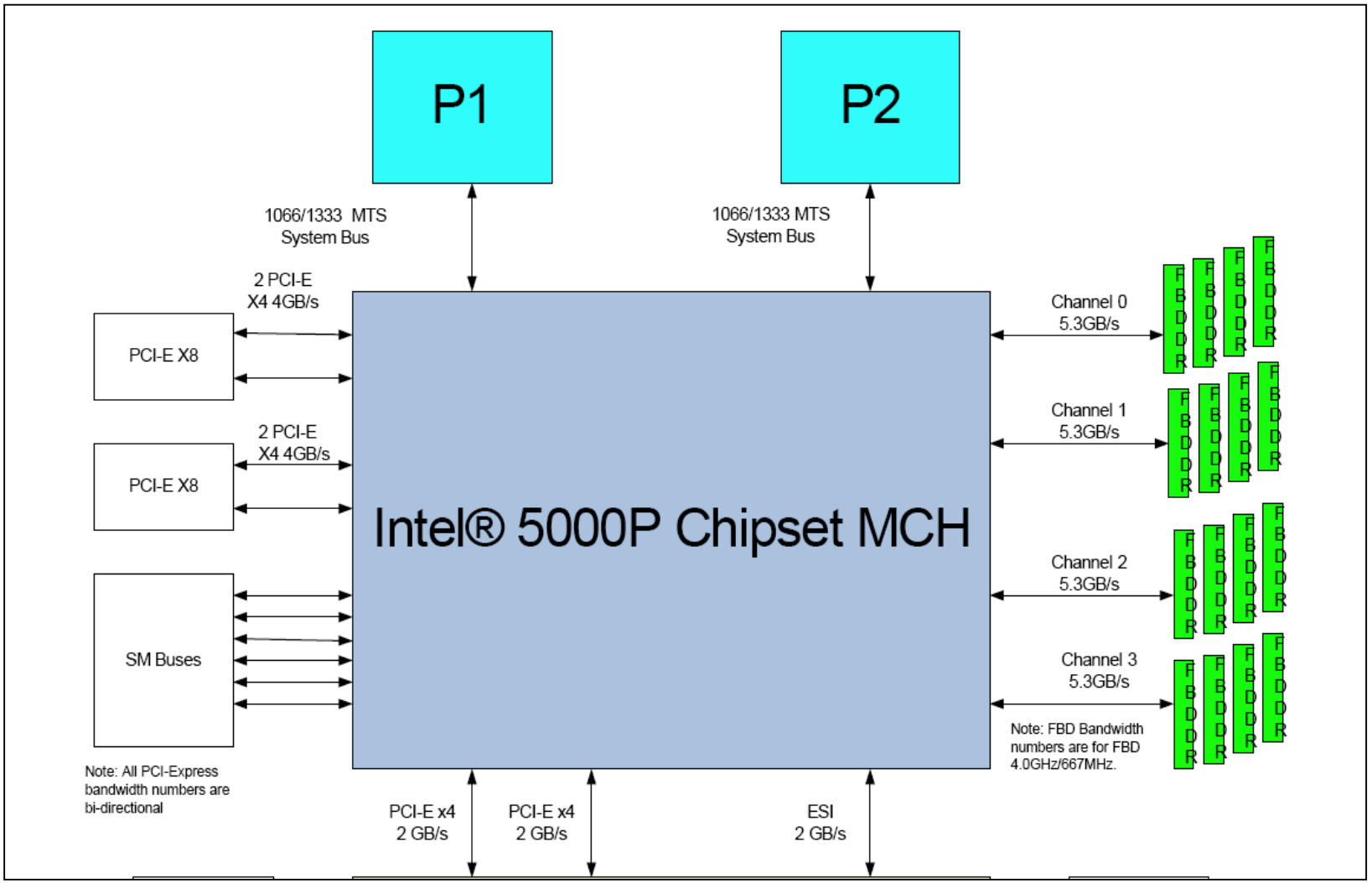
Intel: Clovertown/Harpertown

(don't ask me to explain Intel nomenclature!)

- Pseudo quad-core
  - Two dual-core die MCM
- Clovertown
  - 65 nm process
  - Core 2 vs Core 2 Duo?
- Harpertown
  - 45 nm process
  - New microarchitecture Core 2 Duo Extreme?
- 4 FLOPs/cycle/core



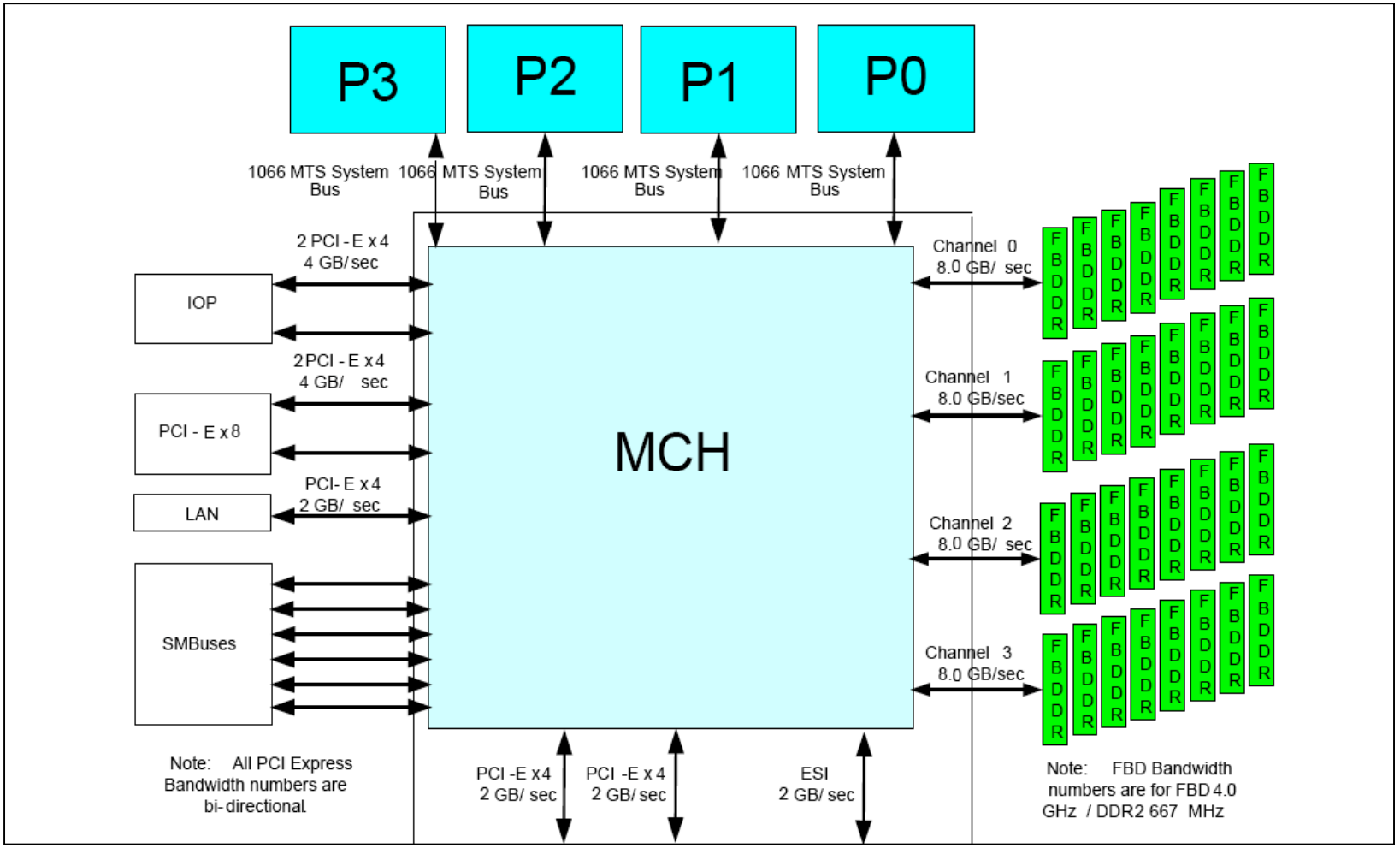
Intel 5000P Northbridge/Memory Controller



NNSA National Nuclear Security Administration

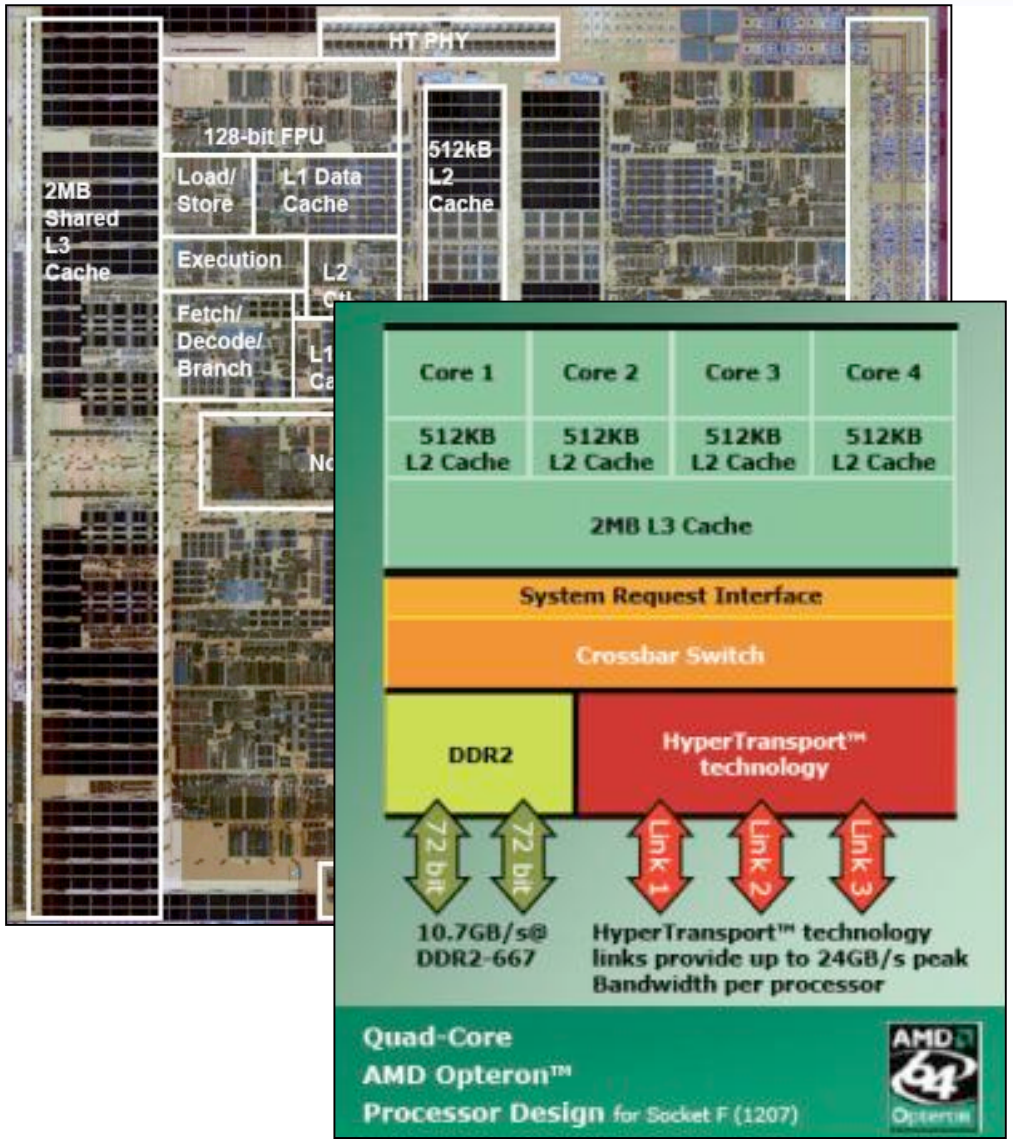
Sandia National Laboratories

Intel 7300 Northbridge: Four Sockets



AMD: Barcelona

- "True" quad-core
- Integrated memory controller
- Uses DDR2/3
  - I.e. no FBDIMM
- Hypertransport for multi-socket nodes
  - NUMA issues
- Dual-channel DDR memory controller
- 2 MB shared LLC
- 4 FLOPs/cycle/core



Sun UltraSPARC Niagra/T2

- "True" eight core die
- Each core supports 8 threads - 64 total threads
- Simple core microarchitecture
- Four dual-channel FBDIMM memory controllers!
- 4 MB shared LLC (L2)
  - 1 MB/MCU
- 1 FLOP/cycle/core!

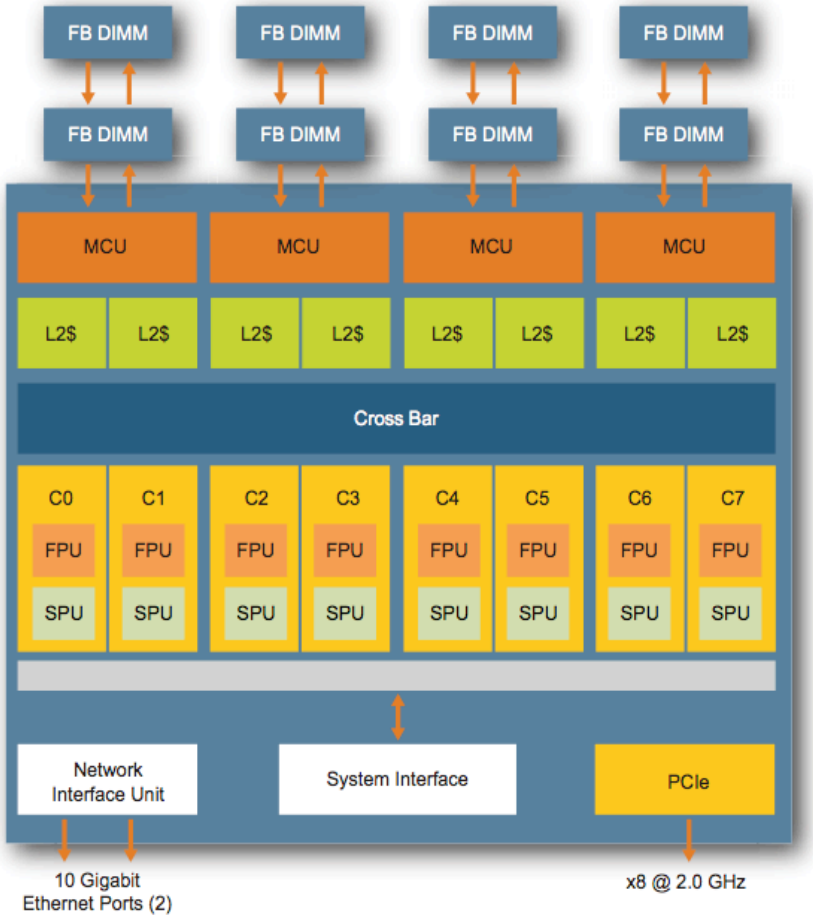


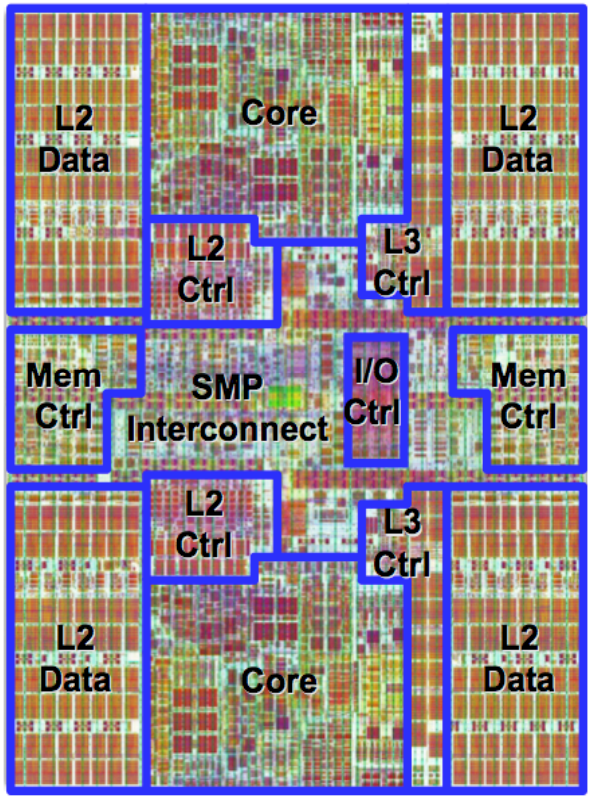
Figure 5. The UltraSPARC T2 processor combines eight cores, memory management, cryptographic support, 10 Gb Ethernet, and PCI Express on a single chip

IBM Power6

- Dual-core!
- High clock rate!
  - Target 4+ GHz
- Dual-channel DDR2/3
- 32 MB LLC (L3)
  - But external to die
- 4 FLOPs/cycle/core

POWER6™ Chip Overview

- **Ultra-high frequency dual-core chip**
  - 7-way superscalar, 2-way SMT core
  - 9 execution units
    - 2LS, 2FP, 2FX, 1BR, 1VMX,1DFU
  - 790M transistors
  - Up to 64-core SMP systems
  - 2x4MB on-chip L2
  - 32MB On-chip L3 directory and controller
  - Two memory controllers on-chip
- **Technology**
  - CMOS 65nm lithography, SOI
- **High-speed elastic bus interface at 2:1 freq**
  - I/Os: 1953 signal, 5399 Power/Gnd

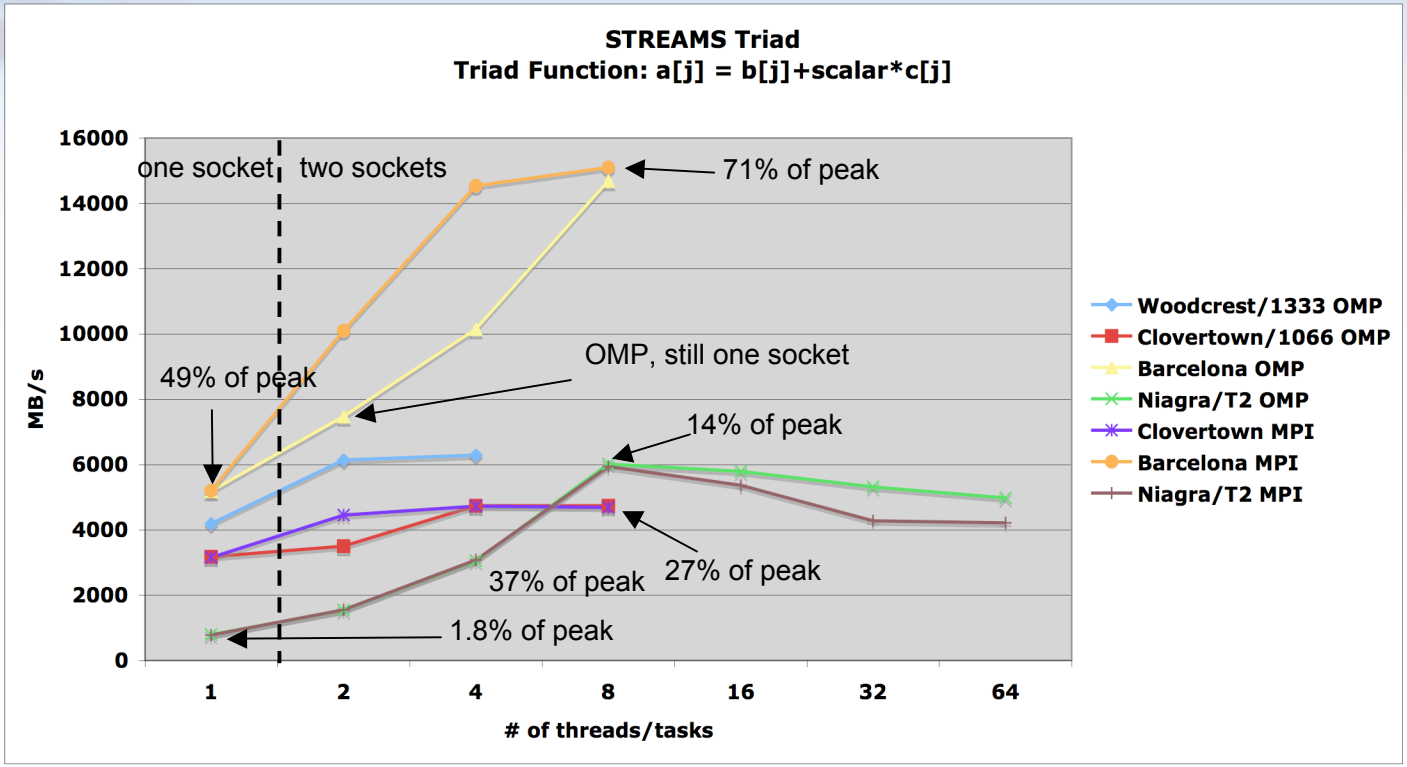


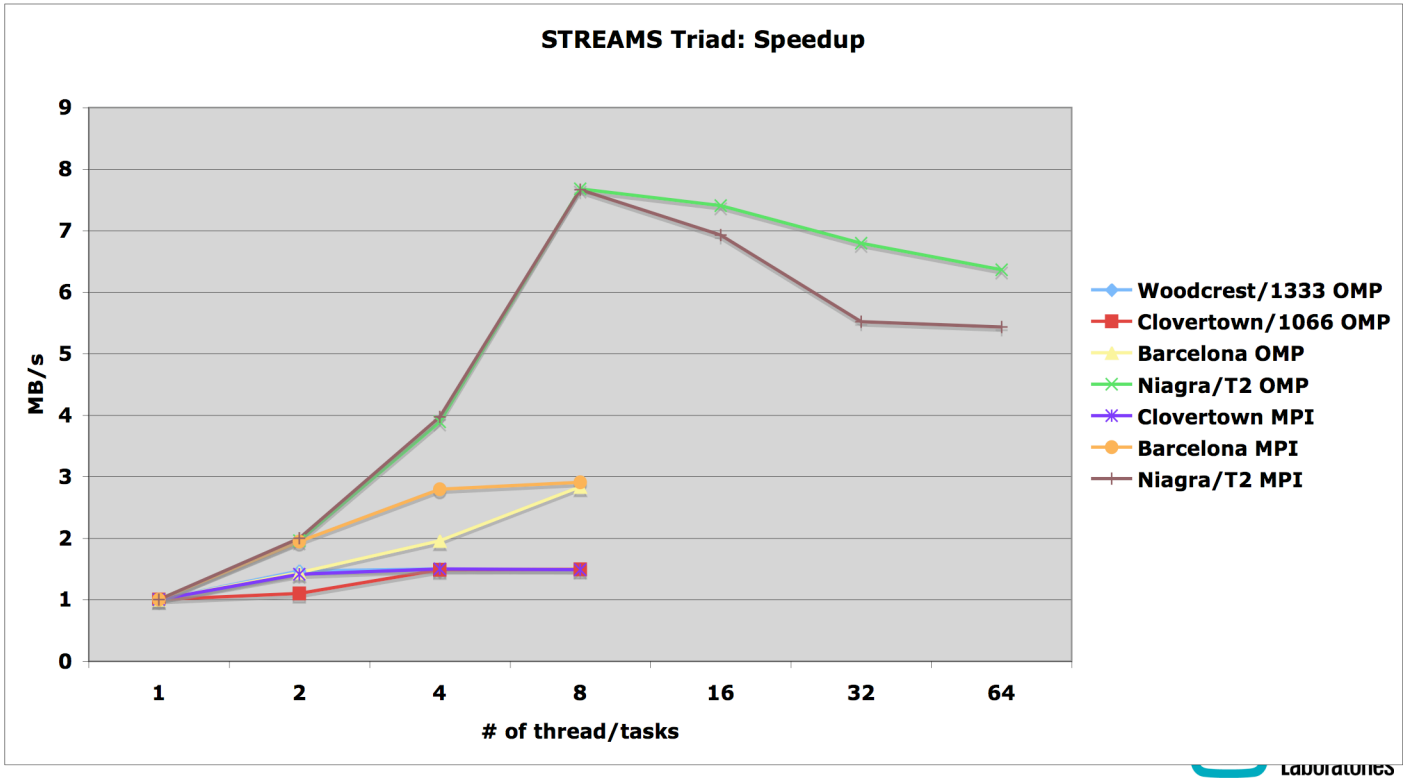
Test Systems by the Numbers

	Clovertown	Barcelona	Niagra/T2
Core frequency	1.86	2.2	1.4
# cores/socket	4	4	8
Execution arch.	out of order	out of order	in order
Total # cores	8	8	8
# sockets	2	2	1
Total Threads	8	8	64
L1 Dcache	32KB/core	64KB/core	8KB/core
L2 cache	2 x 4MB	512KB/core	4 x 1MB
L3 cache	-	2MB	-
DRAM	667MHz FB	667MHz DDR	667MHz FB
# mem. channels	4	2 x 2	8
Peak read BW	21.3	2 x 10.7	42.7
Peak write BW	10.7	same bus	21.3
FLOPs/cycle	4	4	1
Peak FLOPs	59.5	70.4	11.2

STREAMS

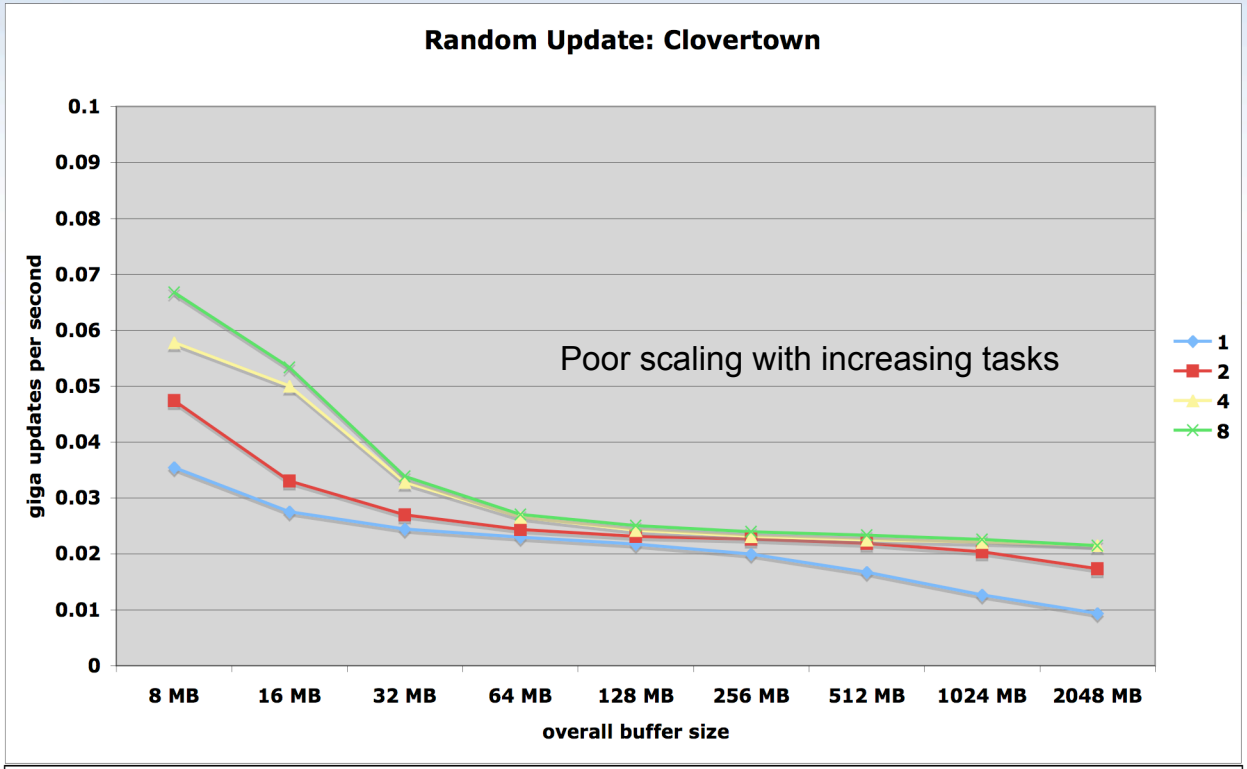
- MPI and OMP
- Two socket nodes
  - 1: 1 core
  - 2: 1 core/socket
  - 4: 2 cores/socket
  - 8: 4 cores/socket
- Clovertown and Barcelona have same peak BW
- Clovertown effectively saturates at 1 core/socket
- Barcelona effectively saturates at 2 cores/socket
- Niagra/T2 becomes FLOPs bound
  - One socket node
- Note Task scheduling differences between MPI and OMP

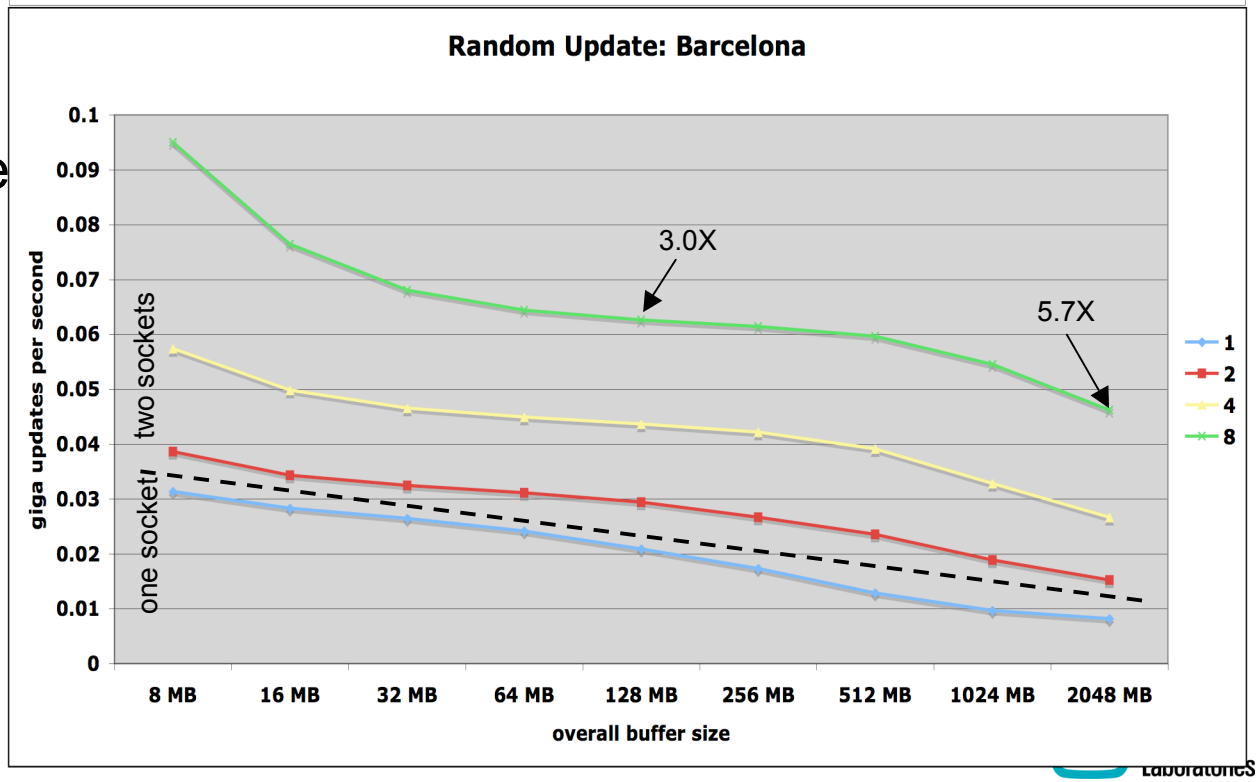




Random Update

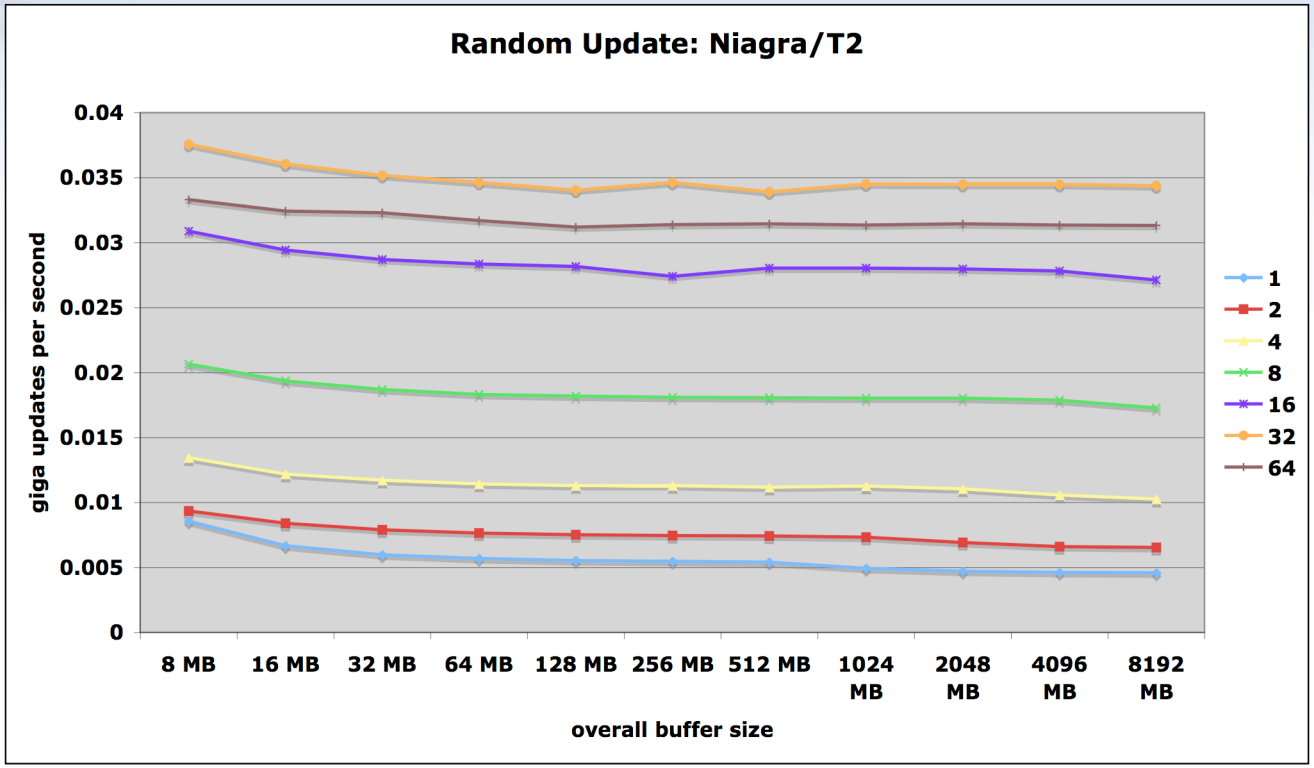
- Two socket nodes
  - 1: 1 core
  - 2: 1 core/socket
  - 4: 2 cores/socket
  - 8: 4 cores/socket
- Varying buffer size
  - Cache effects
  - TLB effects
- Higher is better
- Clovertown: extra cores don't add to performance
- Barcelona: memory subsystem effective at handling extra requests

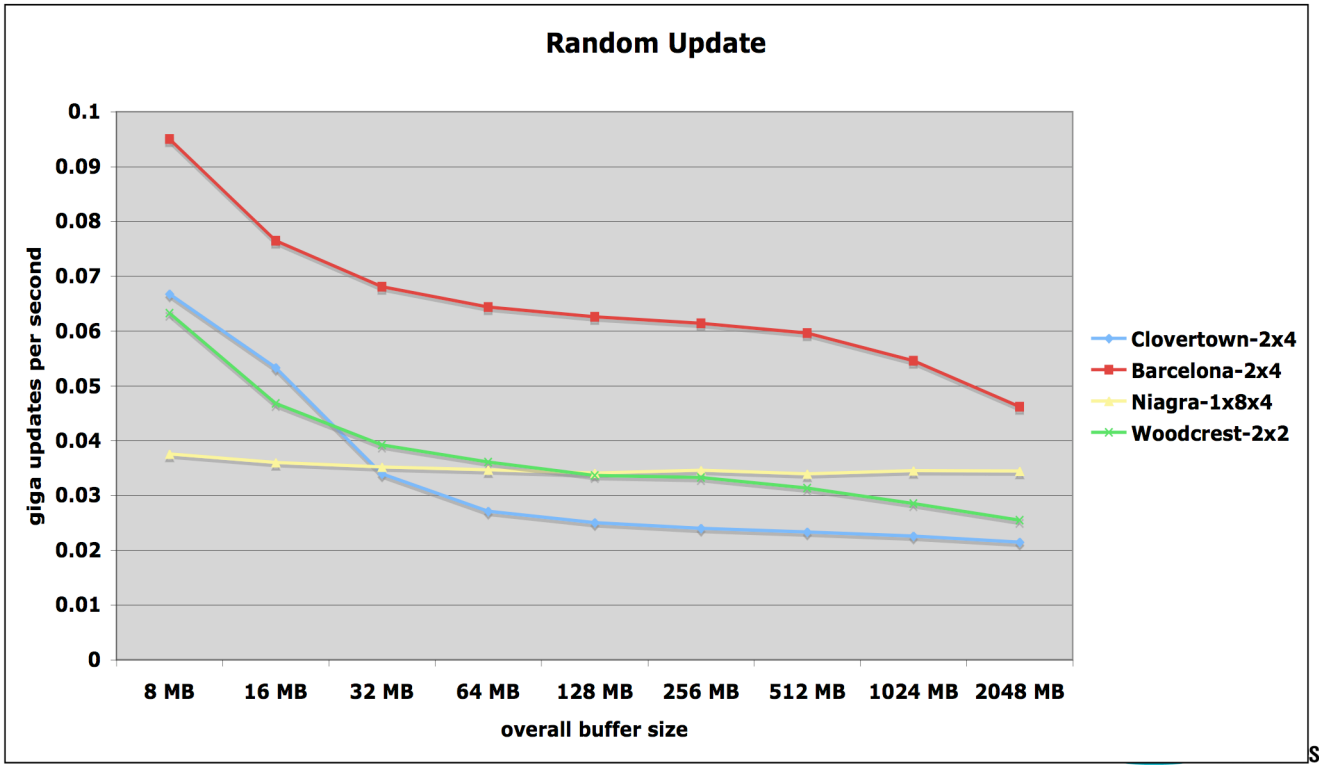




Random Update Cont'd

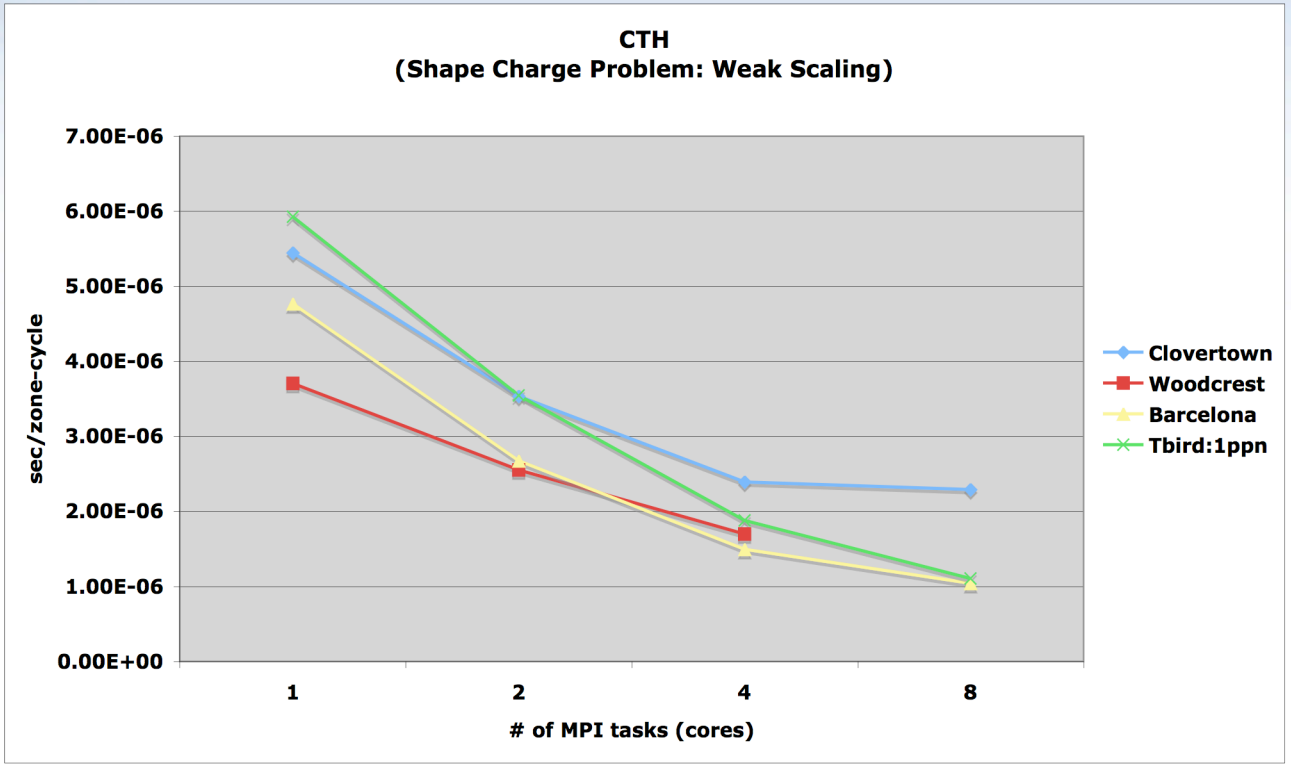
- Niagra/T2 performance
  - Excellent scaling to 16 threads
  - Excellent scaling with buffer size
  - 64 threads < 32 threads
- Best case performance across architectures
  - Barcelona best up to 2048 MB buffer, after that?
  - Intel architecture exhibits lowest performance

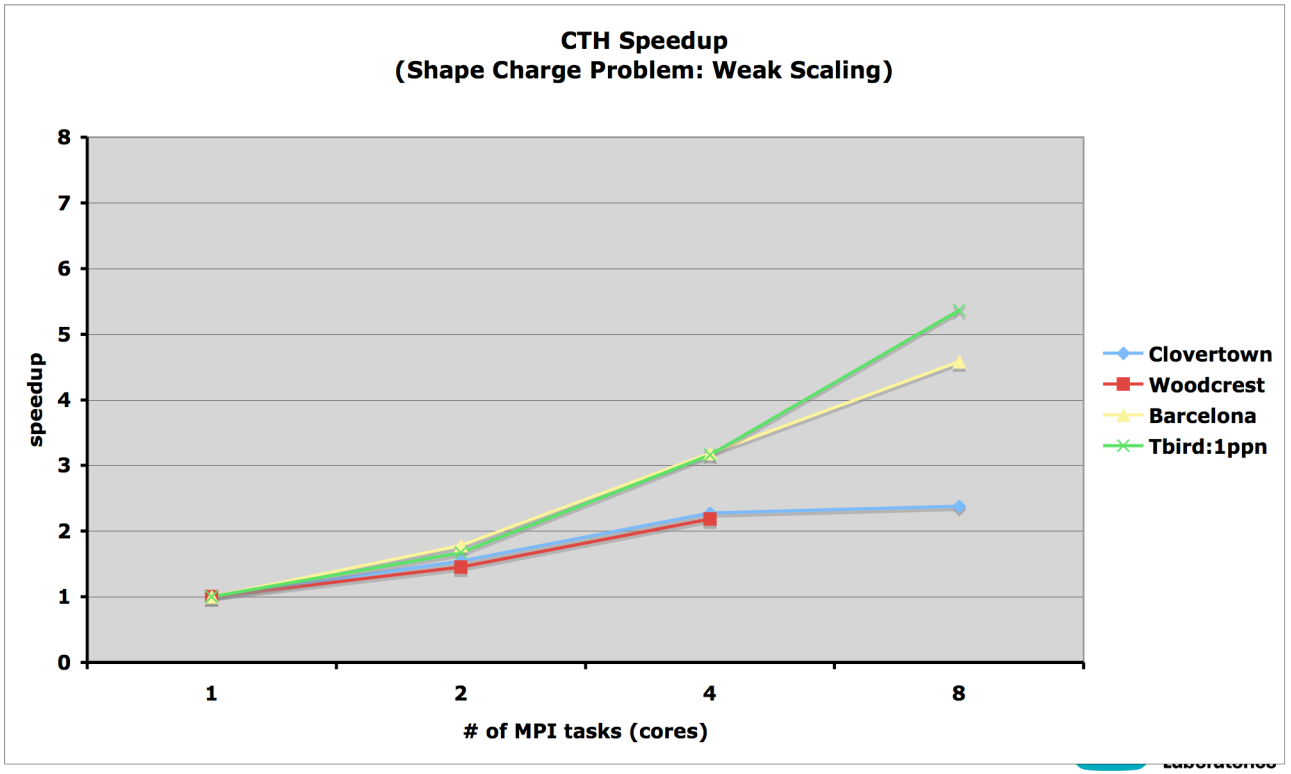




CTH

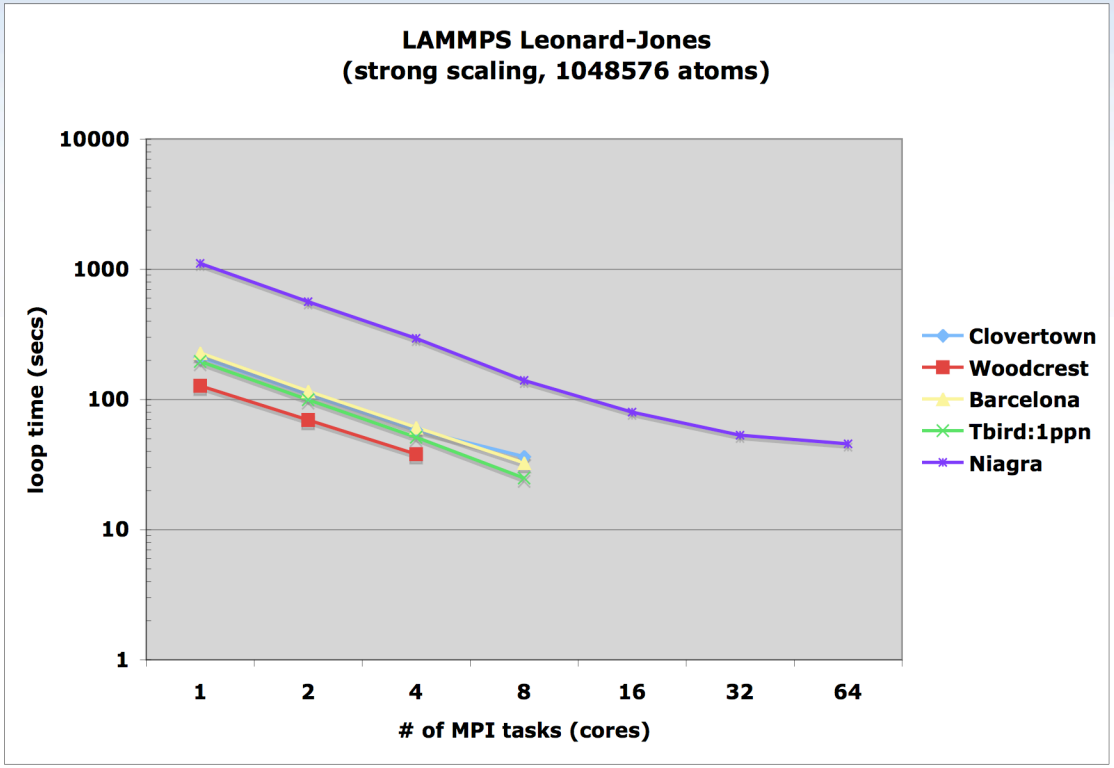
- Shape Charge Problem
- Weak Scaling
  - This is how we plan to use multicore, I.e. spec'ing a minimum GB/core
  - Intracore MPI
- Tbird 1ppn is used as a measure of "ideal" scaling
- Lower is better
- At 8 cores, Barcelona is as fast as 8 Pentium's with Infiniband!
- Clovertown/Woodcrest northbridge issues are evident

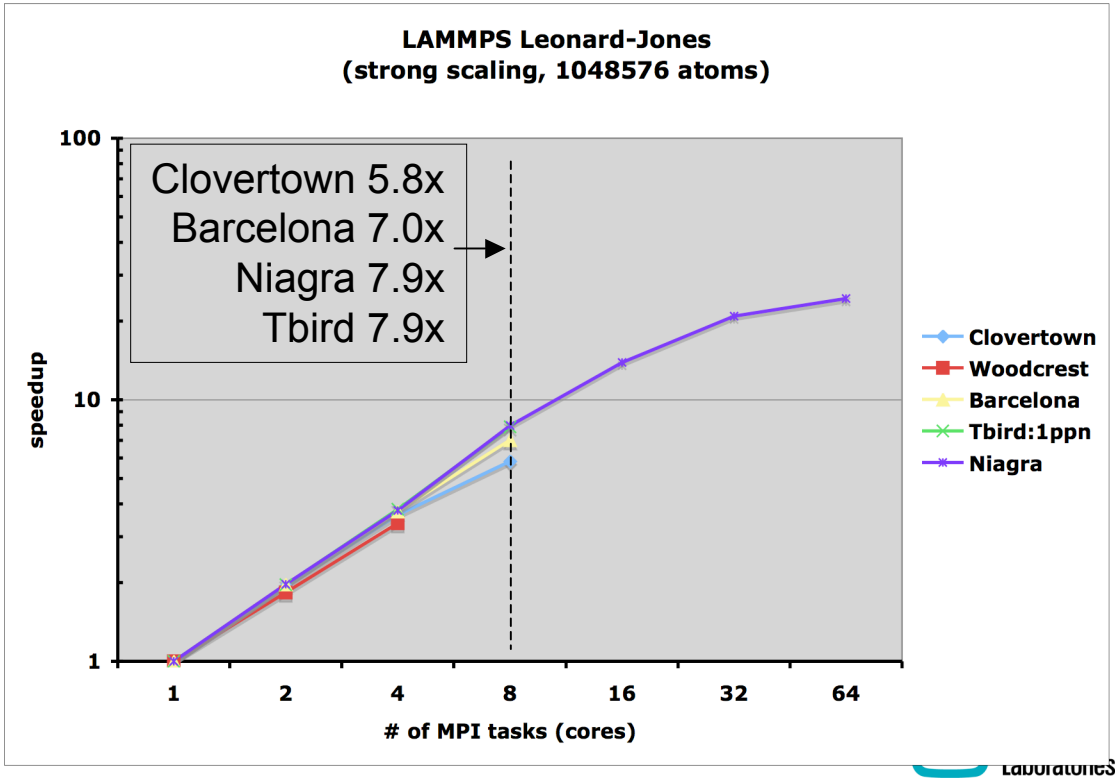




LAMMPS LJ

- Strong Scaling
  - Weak scaling produces similar results
- Lower is better
- All architectures scale very well up to the number of cores/socket
- Niagra never achieves performance of x86-64 architectures, despite excellent scaling





?s & Discussion